

Problem n.2

Prehistoric men crafted stone tools by striking raw stones to obtain the desired shape. Archaeologists rarely find tools, but they often find stone flakes, the waste of the crafting process. The file `stoneflakes.txt` contains the lengths and the widths of stone flakes collected in 75 different archaeological sites.

- a) Identify possible clusters within the data using a hierarchical clustering algorithm (Euclidean distance, Ward linkage). Provide the plot of the dendrogram and qualitatively identify the optimal number of clusters.
- b) Assuming that the clusters identified at point a) have the same covariance structure, formulate a MANOVA model for the geometrical features (width and length) of the stone flakes as a function of the clustering membership. Verify the assumptions of the model. Is there statistical evidence to state that the membership to a cluster has an effect on the mean features of the stone flakes?
- c) Provide confidence intervals for the differences between the mean features of stone flakes belonging to the identified clusters. Use a Bonferroni correction to ensure a 90% global level. Use the computed intervals to comment about the differences among the clusters.